



September 26, 2018

Twitter, Inc.

800 Connecticut Ave NW
Suite 500
Washington, DC 20006

Mr. Peter Micek
Global Policy & Legal Counsel
Access Now
P.O. Box 20429
Greeley Square Station
4 East 27th Street
New York, NY 10001-9998

Dear Mr. Micek,

This letter responds to your invitation for our feedback on the 2018 Ranking Digital Rights Corporate Accountability Index. We are pleased to share some thoughts as we continue to make improvements that further demonstrate our commitment to freedom of expression, privacy, and human rights.

As Access Now noted, Twitter has continued to make significant strides in our transparency efforts over the past year. Beyond the improvements noted in the 2018 Index, we have recently launched our first public comment period to gather feedback from our users about our new dehumanization policy, which prohibits dehumanizing speech against members of protected groups. Additionally, our most recent Transparency Report included differentiation between types of legal requests received, as well as expanded county-level reporting for non-federal requests. Through projects such as these, we continue to hone our efforts to ensure that our users feel safe speaking out on Twitter. We continue to explore new ways to share more information about our work in our transparency report.

With regard to membership in the GNI, we recognize the role and important contribution to policymaking the GNI plays, but are disappointed that the Ranking Digital Rights process continues to take the view that GNI membership is a binary test to demonstrate a human rights commitment. We are currently considering how we can continue to develop our human rights framework and build trust with our customers through greater transparency in this area. To this end, we maintain productive relationships with multi-stakeholder groups including the Center for Democracy & Technology, ICT Watch, the Social Media Exchange, and the Institute for Policy Research and Advocacy, among numerous others who serve on our Trust and Safety Council. We also have numerous partnerships and engagements with human rights and journalistic organizations around the world which are not currently members of the Trust and Safety Council.

More broadly, we are focussed on ensuring we have thoughtful, comprehensive metrics to measure the health of public conversation on Twitter. These metrics are crucial to guiding our work and making progress. Recently, we have partnered with outside experts following a public call for proposals. The details of our partnerships with coalitions led by Leiden and Oxford Universities are included as an appendix to this letter.

We would also like to call attention to the options Twitter has available to protect user privacy. While the majority of Twitter data is public, we do not share nonpublic information without explicit consent, and we have a strong “know your customer” policy when it comes to our data partners. We made a decision to cease supporting Do Not Track (DNT) when it failed to be adopted as an industry standard, which we had hoped it would when Twitter was one of the first companies to adopt it. However we continue to offer our users a range of personalization and data options to control how their information is shared with and used by Twitter. For all intents and purposes, opting out of these settings has the same effect as Do Not Track in protecting user data on Twitter and in some areas goes beyond what DNT provided.

Beyond these personalization and data settings, in April 2018 we released a single, global privacy policy. We worked hard to make this policy transparent about the types of data users share with us, how we use that data, and to highlight the meaningful controls people have over both. Based on feedback we receive, we will continue to refine our privacy policy to make sure it’s as clear and understandable as possible.

Twitter is committed at every level of the company to improving the health of the public conversation, a mission which requires that our users’ privacy and freedom of expression are protected. We look forward to continuing to improve and appreciate your input and feedback.

Sincerely,

A handwritten signature in black ink that reads "Colin Crowell". The signature is written in a cursive style with a large, stylized initial "C".

Colin Crowell
Vice President
Global Public Policy & Philanthropy
Twitter, Inc.

Examining echo chambers and uncivil discourse

Led by Dr. [Rebekah Tromble](#), assistant professor of political science at [Leiden University](#), along with Dr. Michael Meffert at Leiden, Dr. Patricia Rossini and Dr. Jennifer Stromer-Galley at [Syracuse University](#), Dr. Nava Tintarev at [Delft University of Technology](#), and Dr. Dirk Hovy at [Bocconi University](#), this project will develop two sets of metrics: how communities form around political discussions on Twitter, and the challenges that may arise as those discussions develop.

“In the context of growing political polarization, the spread of misinformation, and increases in incivility and intolerance, it is clear that if we are going to effectively evaluate and address some of the most difficult challenges arising on social media, academic researchers and tech companies will need to work together much more closely. This initiative presents an important and promising opportunity for Twitter and our team of researchers to share expertise and work on solutions together,” Tromble said.

The Leiden-led project will primarily focus on two key challenges: echo chambers and uncivil discourse. Based on their past findings, echo chambers, which form when discussions involve only like-minded people and perspectives, can increase hostility and promote resentment toward those not having the same conversation. The project’s first set of metrics will assess the extent to which people acknowledge and engage with diverse viewpoints on Twitter.

The second set of metrics will focus on incivility and intolerance in Twitter conversations. The group has found that while incivility, which breaks norms of politeness, can be problematic, it can also serve important functions in political dialog. In contrast, intolerant discourse — such as hate speech, racism, and xenophobia — is inherently threatening to democracy. The team will therefore work on developing algorithms that distinguish between these two behaviors.

The project brings together scholars with different backgrounds and expertise, bridging political science, communication, and computer science. The six scholars working together across four universities will combine deep social scientific subject expertise with strong technical expertise to develop and adapt measures of healthy conversations on Twitter.

Bridging gaps between communities on Twitter

Professor Miles Hewstone and John Gallacher at [the University of Oxford](#), in partnership with Dr. Marc Heerdink at the [University of Amsterdam](#), will be studying how people use Twitter, and how exposure to a variety of perspectives and backgrounds can decrease prejudice and discrimination.

“We’re very excited about the opportunity to work with Twitter on investigating the important social challenges of a digitally connected world,” said Miles Hewstone, Ph.D., professor of social psychology at Oxford University.

“Evidence from social psychology has shown how communication between people from different backgrounds is one of the best ways to decrease prejudice and discrimination. We’re aiming to investigate how this understanding can be used to measure the health of conversations on Twitter, and whether the effects of positive online interaction carry across to the offline world.”

The project builds on Professor Hewstone's [longstanding work to study intergroup conflict](#). When the communication between groups contains more positive sentiments, cooperative emotions, and more complex thinking and reasoning from multiple perspectives, [prejudice is reduced and relations can improve](#). This was previously demonstrated in a variety of contexts, including in post-conflict community building efforts in [Northern Ireland](#) as well as in [online communities](#). As part of the project, text classifiers for language commonly associated with positive sentiment, cooperative emotionality, and integrative complexity will be adapted to the structure of communication on Twitter.